

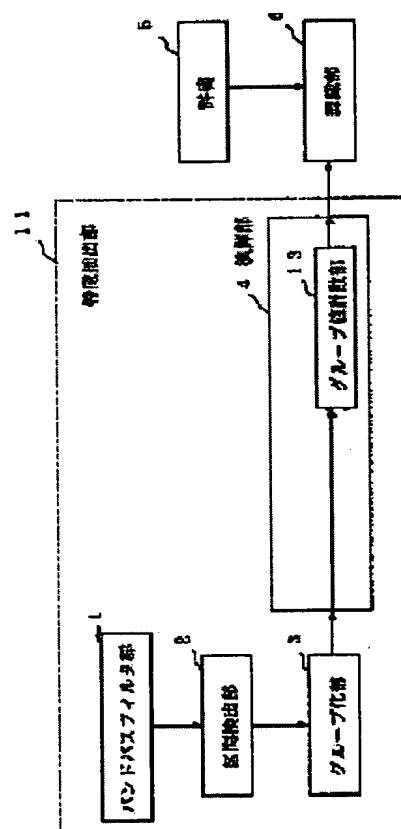
VOICE FEATURE EXTRACTION DEVICE

Patent number: JP5019782
Publication date: 1993-01-29
Inventor: KATO YOSHINAGA
Applicant: RICOH KK
Classification:
- international: (IPC1-7): G10L3/00
- european:
Application number: JP19910232212 19910820
Priority number(s): JP19910232212 19910820; JP19910130641 19910502

Report a data error here

Abstract of JP5019782

PURPOSE: To improve the recognition precision more by obtaining a feature pattern which absorbs the fine variation of a voice.
CONSTITUTION: The band-pass filter group 1 of a feature extraction part 11 samples an input voice at constant intervals of time and takes a frequency analysis. A section detection part 2 detects the section of the voice and extracts only sections corresponding to the voice part from the analytic value which is the result of the frequency analysis. A grouping part 3 puts analytic values extracted by the section detection part 2 together in a group in an analytic axis direction for each dimension to divide them, group by group. An arithmetic part 4 calculates against the analytical result of the group to generate the new feature pattern.

Data supplied from the **esp@cenet** database - Worldwide

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平5-19782

(43) 公開日 平成5年(1993)1月29日

(51) Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/00	5 2 1 D	8622-5H		
	K	8622-5H		

審査請求 未請求 請求項の数6(全21頁)

(21) 出願番号 特願平3-232212

(22) 出願日 平成3年(1991)8月20日

(31) 優先権主張番号 特願平3-130641

(32) 優先日 平3(1991)5月2日

(33) 優先権主張国 日本 (J P)

(71) 出願人 000006747

株式会社リコー

東京都大田区中馬込1丁目3番6号

(72) 発明者 加藤 喜永

東京都大田区中馬込1丁目3番6号 株式

会社リコー内

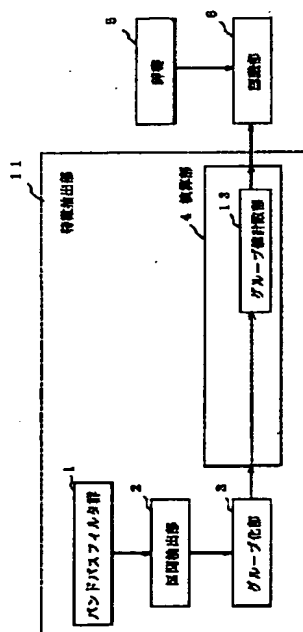
(74) 代理人 弁理士 植本 雅治

(54) 【発明の名称】 音声特徴抽出装置

(57) 【要約】

【目的】 音声の微小な変動を吸収した特徴パターンを得ることによって認識精度を一層向上させることが可能である。

【構成】 特徴抽出部11のバンドパスフィルタ群1では、入力音声を一時間ごとにサンプリングし周波数分析する。区間検出部2では、音声の区間を検出し、周波数分析した結果の分析値から音声部分に相当するものだけを取り出す。グループ化部3では、区間検出部2で取り出された分析値を分析軸方向にいくつかの次元ごとにグループとしてまとめ上げて、グループに分割する。演算部4では、グループ内の分析値に対して演算を施し、新たな特徴パターンを生成する。



(2)

特開平5-19782

1

2

【特許請求の範囲】

【請求項1】 音声認識を行なわせるため入力音声から特徴パターンを抽出する音声特徴抽出装置において、入力音声を音響分析した結果の各分析値をグループに分割するグループ化手段と、分割された各グループ内の分析値に所定の演算処理を施して特徴量を生成する演算処理手段とを備えていることを特徴とする音声特徴抽出装置。

【請求項2】 請求項1記載の音声特徴抽出装置において、分割された各グループは、互いに隣接したグループ間において少なくとも一部が重複している分析値を有していることを特徴とする音声特徴抽出装置。

【請求項3】 請求項1または2記載の音声特徴抽出装置において、前記各グループは、分析軸方向に分割されるようになっていることを特徴とする音声特徴抽出装置。

【請求項4】 請求項1, 2, または3記載の音声特徴抽出装置において、前記各グループは、時系列方向に分割されるようになっていることを特徴とする音声特徴抽出装置。

【請求項5】 請求項1, 2, 3, または4記載の音声特徴抽出装置において、前記演算手段は、各グループ内の分析値を加算し特徴量を生成するようになっていることを特徴とする音声特徴抽出装置。

【請求項6】 請求項1, 2, 3, または4記載の音声特徴抽出装置において、各グループ内の分析値に所定の演算処理を施して生成された特徴量に対し正規化処理を施し、最終的な特徴ベクトルを生成するようになっていることを特徴とする音声特徴抽出装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、音声認識システムにおいて、入力音声からその特徴パターンを抽出する音声特徴抽出装置に関する。

【0002】

【従来の技術】 一般に、音声認識システムは、図30に示すように、特徴抽出部51と、辞書52と、識別部53とから構成されている。このような構成では、音声が入力すると、先づこの入力音声から音声認識に必要な特徴パターンを特徴抽出部51で抽出する。具体的には、入力音声に対し音響分析を施し、入力音声を一定時間ごとにサンプリングし、周波数分析して10～20次元のベクトルの時系列として表現する。

【0003】 音響分析された結果のこのようなベクトルの時系列データ（分析値）は、音声区間検出等の処理が施された後、特徴パターンとして用いられ、識別部53では、このようにして得られた特徴パターンを辞書52内の種々の特徴パターンと比較し、既知のカテゴリのいずれかに分類して、識別結果として出力する。

【0004】 ところで、人間の発する音声は不安定であ

り、例えば同一人が同じ言葉を発声したとしても、音声長のみならず、周波数成分の分布、すなわち分析軸方向の分布も、その都度、変化することが知られている。

【0005】 図31(a)は入力音声を周波数分析した結果のパターンの一例を示す図であり、図31(b)は入力音声図31(a)の場合に比べて微小変動したときの周波数分析結果のパターンの一例を示す図である。なお、図31(a), (b)の例では、入力音声を1フレーム当たり15次元で分析している。

【0006】 図31(a), (b)に示すような各パターンは、本来同一のカテゴリに分類されるべきものであるが、同一人が同じ音声を発声する場合にも、発声者が声の高さを変えたりするなどの微小変動によって互いに明らかに異なったものとなり、図31(a), (b)の各パターンをマッチング手法などを用いて認識させると、両パターンはそれぞれ異なるカテゴリに分類され、互いに異なった音声のものとして認識されてしまう。

【0007】 パターンの局所的な変動の影響を避けるために、従来では、音響分析の結果得られた分析値を特徴パターンとしてそのまま用いるのではなく、時系列方向に数フレームの窓を設け、その窓内に含まれる分析値を1つの特徴ベクトルとし、その窓を時系列方向に1フレームずつシフトすることによって、新たな特徴パターンを作成する方式が提案されている（文献「LVQ-HMM音韻認識と予測LRパーザを用いた文節認識 信学技報, SP89-100」参照）。

【0008】 図32はこの文献に開示の手法による特徴ベクトルの作成の仕方を示す図である。なお、図32では、分析次元を“15”，窓のフレーム数を“3”としている。この仕方では、分析値を時系列方向に数フレームごとにまとめて新たな特徴パターンとすることにより、時系列方向の局所的変動を吸収することができる。

【0009】

【発明が解決しようとする課題】 しかしながら、上述した従来の音声特徴抽出装置では、時系列方向の局所的変動を吸収することはできるものの、分析軸方向の変動についてはこれを吸収することができず、認識精度を向上させるには限度があった。

【0010】 本発明は、上記従来技術の欠点を補い、音声の微小な変動を吸収した特徴パターンを得ることによって認識精度を一層向上させることの可能な音声特徴抽出装置を提供することを目的としている。

【0011】

【課題を解決するための手段】 上記目的を達成するために請求項1記載の発明は、入力音声を音響分析した結果の各分析値をグループに分割するグループ化手段と、分割された各グループ内の分析値に所定の演算処理を施して特徴量を生成する演算処理手段とを備えていることを特徴としている。

【0012】 また、請求項2記載の発明は、分割された

(3)

特開平5-19782

3

各グループが、互いに隣接したグループ間において少なくとも一部が重複している分析値を有していることを特徴としている。

【0013】また、請求項3記載の発明は、前記各グループが、分析軸方向に分割されるようになっていることを特徴としている。

【0014】また、請求項4記載の発明は、前記各グループが、時系列方向に分割されるようになっていることを特徴としている。

【0015】また、請求項5記載の発明は、前記演算手段が、各グループ内の分析値を加算し特徴量を生成するようになっていることを特徴としている。

【0016】また、請求項6記載の発明は、各グループ内の分析値に所定の演算処理を施して生成された特徴量に対し正規化処理を施し、最終的な特徴ベクトルを生成するようになっていることを特徴としている。

【0017】

【作用】請求項1記載の発明では、入力音声を音響分析した結果の各分析値をグループに分割し、分割された各グループ内の分析値に所定の演算処理を施して特徴量を生成することにより、音声の微小変動を吸収することができる。

【0018】また、請求項2記載の発明では、分割された各グループは、互いに隣接したグループ間において少なくとも一部が重複している分析値を有しているため、グループ間にまたがる音声の微小変動をも吸収することができる。

【0019】また、請求項3記載の発明では、分析軸方向に分割されるようになっているので、分析軸方向の微小変動を吸収することができる。

【0020】また、請求項4記載の発明では、時系列方向に分割されるようになっているので、時系列方向の微小変動を吸収することができる。

【0021】また、請求項5記載の発明では、各グループ内の分析値を加算し特徴量を生成するようになっているので、特徴量を削減することができ、辞書とのマッチングにおける計算量を減らすことができる。

【0022】また、請求項6記載の発明では、各グループ内の分析値に所定の演算処理を施して生成された特徴量に対し正規化処理を施し、最終的な特徴ベクトルを生成するようになっているので、音量の大小差を吸収することができる。

【0023】

【実施例】以下、本発明の一実施例を図面に基づいて説明する。図1は本発明の第1の実施例のブロック図である。図1を参照すると、この第1の実施例では、入力音声に基づき特徴パターンを生成する特徴抽出部11と、種々の標準パターンが予め格納されている辞書5と、特徴抽出部11で得られた特徴パターンと辞書5に予め格納されている種々の標準パターンとをDPマッチング等

4

により照合し、認識結果を得る認識部6とが設けられている。

【0024】特徴抽出部11は、入力音声を一時間ごとサンプリングし周波数分析するバンドパスフィルタ群1と、音声の区間を検出し、周波数分析した結果の分析値、すなわちパターンから音声部分に相当するものだけを取り出す区間検出部2と、区間検出部2で取り出された分析値を分析軸方向にいくつかの次元ごとにグループとしてまとめ上げて、グループに分割するグループ化部3と、グループ内の分析値に対して所定の演算を施し、新たな特徴パターンを生成する演算部4とを有している。第1の実施例においては、演算部4は、グループ内の要素、すなわち分析値の和を計数するグループ値計数部13により構成されている。

【0025】次にこのような構成における第1の実施例の音声特徴抽出装置の動作を図2のフローチャートを用いて説明する。なお、以下では、音響分析時の総フレーム数をMとし、また分析軸方向のグループ数をnとし、1つのグループ内の分析値の個数をqとする。また、この第1の実施例では、1つのフレーム内の各分析値に対し、第1番目の分析値から第q番目の分析値までを1つのグループとして分割し、第(q+1)番目の分析値から第(2q)番目の分析値までを次のグループとして分割するようになっている。従って、各分析値は、分析軸方向に重複せずにグループ分割される。

【0026】先づ、ステップS1では、フレームの番号jを“1”に初期設定し、ステップS2では、グループの番号iを“1”に初期設定する。次いで、ステップS3では、グループ内の分析値f_{ijk}に対して所定の演算処理を施す。なお、ここでkは1つのグループ内の各分析値を示すサフィックスであり、1つのグループ内にはq個の分析値があるので、kは1～qの値をとる。図2の例では、この演算処理は、グループ内の各分析値f_{ijk}の合計をとる加算演算となっており、この場合に、演算の結果、j番目のフレーム、i番目のグループにおける各分析値f_{ijk}の合計値A_{ji}が新たな特徴量として得られる。いまの場合、j、iはそれぞれ“1”となっているので、1番目のフレーム、1番目のグループに対し、新たな特徴量A₁₁が得られる。

【0027】次いで、j番目のフレームの次のグループに対しても同様の処理を行なうため、iを“1”だけ歩進し、再びステップS3に戻る。これにより、1番目のフレーム、2番目のグループに対し、新たな特徴量A₁₂が得られる。このような処理を繰り返行ない、1番目のフレームのn番目のグループ(最後のグループ)に対して、新たな特徴量A_{1n}が得られる。A_{1n}が得られた後、iはnよりも大きくなるので(ステップS5)、ステップS6に進み、jを“1”だけ歩進し、再びステップS2に戻る。

【0028】これにより、次のフレーム、いまの場合、

(4)

特開平5-19782

5

2番目のフレームの各グループについて1番目のフレームにおいてなされたと同様の処理がなされ、新たな特徴量として、 $A_{21} \sim A_{2n}$ が得られる。このような処理を順次に繰り返して、 M 番目のフレーム（最後のフレーム）の n 番目のグループ（最後のグループ）に対して、新たな特徴量 A_{Mn} が得られると、次に j が M よりも大きくなるので（ステップS7）、全ての演算処理を終了する。

【0029】以上のようにして M フレーム数分の n 次元特徴ベクトルからなる新たな特徴パターン $A_{11} \sim A_{Mn}$ を得ることができる。図3は j 番目のフレームについてなされる上記グループ分割処理並びに演算処理の具体例を示す図である。この例では、バンドパスフィルタ群1が15個のバンドパスフィルタからなり、15個のバンドパスフィルタによって1フレーム当たり分析次元が“15”の分析値（すなわち15個の分析値）が得られたとした場合が示されている。この際に、1フレーム当りのグループ数 n を“3”にし、1つのグループ内の分析値の個数 q を“5”にすると、1つのフレーム（例えば j 番目のフレーム）では、分析軸方向に5つごとに1つのグループ化がなされ、各グループ（3つのグループ）においてグループ内の5個の分析値の合計をとることにより、1つのフレーム j で A_{j1} , A_{j2} , A_{j3} の3つの特徴量を1つの特徴ベクトルとして得ることができる。

【0030】このように第1の実施例では、音響分析された結果の分析値に対し、これを分析軸方向にグループ化し、各グループ内で所定の演算処理を施しているの、その結果、各グループにおける音声の分析軸方向の微小変動を吸収した新しい特徴パターン $A_{11} \sim A_{Mn}$ を得ることができた。

【0031】ところで、上述した第1の実施例において得られる新しい特徴パターン $A_{11} \sim A_{Mn}$ をそのまま用いると、音声の大小などにより分析値が変動し、分析軸に沿った分布の形状が同じであって本来同一カテゴリに分類されるべきパターンであっても各値が異なるために正しく分類できなくなる可能性が生ずる。

【0032】図4(a), (b)は分析軸に沿った分布の形状が互いに同じであるが、各値が異なっている特徴ベクトルの一例を示す図であり、図4(b)の特徴ベクトルは図4(a)の特徴ベクトルに比べて2倍の音量のものとなっている。第1の実施例では、図4(a), (b)のような特徴ベクトルに対し、それぞれ異なったカテゴリのものに分類してしまう恐れがある。

【0033】図5は本発明の第2の実施例のブロック図であって、この第2の実施例では、第1の実施例をさらに改良し、第1の実施例における上記のような問題を回避するようにしている。すなわち、この第2の実施例の特徴抽出部12では、演算部14が、グループ値計数部13と、正規化部15とから構成され、第1の実施例の演算部4における演算処理機能に加えて、さらに正規化処理機能を有しており、第1の実施例の演算処理によ

6

て得られた時系列フレームの特徴パターン $A_{11} \sim A_{Mn}$ の所定フレーム数分の各特徴量を正規化し最終的な特徴パターンとするようになっている。

【0034】次にこのような第2の実施例の音声特徴抽出装置の処理動作を図6のフローチャートを用いて説明する。演算部14は、第1の実施例の演算部4と同様に、先づ図2のフローチャートに示す処理を行ない、特徴パターン $A_{11} \sim A_{Mn}$ を得た後、図6のフローチャートに示す処理を行なう。図6の処理では、ステップS11でフレームの番号 j を“0”に初期設定し、次いで、ステップS12で正規化処理を行なう。

【0035】この正規化処理では、正規化用の窓の大きさが予め設定されている必要がある。この窓を（フレーム数 m ） \times （分析軸方向次元数 n ）の大きさに設定すると、先づ、この窓に含まれる特徴量 A_{ri} （ $i=1 \sim n$, $r=j+1 \sim j+m$ ）を全て合計し、合計値 S を求める。次いで、この窓に含まれる各特徴量 A_{ri} を合計値で除算し正規化して、最終的な特徴量 $B(j+1)h$ として求める。ここで、 h は $1 \sim (m \times n)$ までの値をとるサフィックスであり、このことからわかるように、この処理では、正規化とともに m 行 n 列のマトリックスからなる窓内の特徴量を1行（ $m \times n$ ）列の特徴ベクトルに変換している。

【0036】いまの場合、 j が“0”に初期設定されているので、窓は j が“1” \sim “ m ”の範囲の設定され、この範囲のフレームに含まれている特徴量 A_{ri} に正規化処理がなされ、これらの特徴量は（ $m \times n$ ）列の特徴ベクトルに変換される。

【0037】しかる後、ステップS13において j を“1”だけ歩進し、再びステップS12に戻る。この結果、窓は、時系列方向に“1”だけずらされ、ずらされた窓に含まれる各特徴量 A_{ri} に対しても、上記と同様の正規化処理がなされ、1行（ $m \times n$ ）列の特徴ベクトルに変換される。

【0038】このようにして、窓を時系列方向に順次にずらしながら、その窓に含まれる各特徴量に上述した処理を施し、 j が“ $M-m-1$ ”となり、最後の窓についての処理がなされると、最終的な特徴パターンとして、時系列数が（ $M-m-1$ ）、ベクトル次元数が（ $m \times n$ ）の特徴パターン $B_{11} \sim B_{(M-m-1)(m \times n)}$ を得ることができる。

【0039】しかる後、ステップS13において j が“1”だけ歩進され、ステップS14において、 j が（ $M-m$ ）となると、全ての演算処理を終了する。

【0040】図7は第2の実施例における演算処理の具体例を示す図であり、第1の実施例の説明において用いた図3に対応している。すなわち、図7の例では、図3における処理がなされ、特徴パターン A_{j1} , A_{j2} , A_{j3} が求められた後、これに正規化処理を施す場合が示されている。この際、窓 W を規定するフレーム数 m を“3”

(5)

特開平5-19782

7

とし、窓の大きさを“ 3×3 ”のものにして、この窓Wを時系列方向に順次にずらしながら正規化処理を行なう。窓Wが実線で示す位置にあるときには、この位置に含まれる特徴量 $A_{j1} \sim A_{(j+2)3}$ に対して正規化処理を行ない、9次元の1つの特徴ベクトル $B_{j1} \sim B_{j9}$ を得る。次いで、窓Wが破線で示す位置に移動するときには、この位置に含まれる特徴量 $A_{(j+1)1} \sim A_{(j+3)3}$ に対し同様の正規化処理を行ない、9次元の1つの特徴ベクトル $B_{(j+1)1} \sim B_{(j+1)9}$ を得ることができる。

【0041】これによって、図4(a)、(b)のような特徴ベクトルは、最終的には図8(a)、(b)のような正規化された特徴ベクトルにそれぞれ変換される。図8(a)と図8(b)とを比べればわかるように、正規化処理によって、図4(a)、図4(b)の特徴ベクトルは、最終的に、分布軸に沿った分布の形状が互いに同じであるとともに、音量の大小差が吸収されてその値が互いに同じものとなる。この結果、当初各分析値が異なっているにもかかわらず同一カテゴリに分類されるべきパターンを、同一のカテゴリに正しく分類し認識能力を向上させることができる。

【0042】なお、上述の各実施例において、分析軸方向のグループ数 n は、任意所望のものに可変に設定することができ、これに伴って、1つのグループ内の分析値の個数 q も可変に設定できる。また、第2の実施例において、窓を1フレームづつずらすとしたが、例えば2フレームづつずらすようにしても良く、ずらす幅は任意所望のものに可変に設定できる。但し、ずらす幅を大きくすることにより、これに応じて、特徴ベクトルの系列数は減少する。

【0043】このように、第1、第2の実施例では、グループ内における音声の分析軸方向の微小変動を吸収することはできるが、グループ間で共有する要素、すなわち共有の分析値がないため、グループ間にまたがる微小変動については吸収できない。

【0044】図9は本発明の第3の実施例のブロック図であり、この第3の実施例では、第1、第2の実施例における上記問題を解決することを意図している。すなわち、図9を参照すると、この第3の実施例の特徴抽出部16は第1の実施例の特徴抽出部11と対応した構成となっており、入力音声を音響分析した結果の各分析値を分析軸方向にグループに分割するグループ化部17と、分割された各グループ内の分析値に所定の演算処理（加算処理）を施して特徴パターンを生成する演算部18とを有しているが、第3の実施例では、グループ化部17が各分析値を分析軸方向に一部重複させながらグループ分割するようにしている。すなわち、各要素（各分析値）を例えば1つづつずらしながらグループ分割するようにしている。

【0045】また、演算部18は、第1の実施例と同様のグループ値計数部19により構成されており、グルー

8

プ値計数部19は、上記のように一部重複させて分割されたグループ内の要素、すなわち分析値の和を計数するようになっている。

【0046】次にこのような構成における第3の実施例の音声特徴抽出装置の動作を図10のフローチャートを用いて説明する。なお、以下では、音響分析時の総フレーム数を M とし、音響分析次元数（1つのフレーム内の分析値の総数）を k とし、1つのグループ内の分析値の個数を q とする。

【0047】先づ、ステップS21では、1つのフレーム内における分析値の番号 i を“1”に初期設定し、ステップS2では、1つのグループ内の要素番号、すなわち分析値の番号 x を“1”に初期設定する。次いで、ステップS23～S24では、分析値の番号が $(i+x+1) \sim (q+1)$ の範囲を1つのグループとして特定し、このグループ内の分析値 f_k に対して所定の演算処理を施す。なお、ここで k は1つのグループ内の各分析値を示すサフィックスであり、1つのグループ内には q 個の分析値があるので、 k は最初、 i の値をとり、最後は、 $(i+q-1)$ の値をとる。図10の例では、この演算処理は、1つのグループ内の各分析値 k の合計をとる加算演算となっており、この場合に、演算の結果、1番目の分析値から $(i+q-1)$ 番目の分析値までの1つのグループにおける各分析値 f_k の合計値 A_i が新たな特徴量として得られる。いまの場合、 i は“1”となっているので、1つのフレームの1番目の分析値から q 番目の分析値までが最初のグループとして分割され、このグループに対し、新たな特徴量 A_1 が得られる。

【0048】次いで、 i を“1”だけ歩進し（ステップS26）、再びステップS22に戻る。これにより、同様の演算により、2番目の分析値から $(q+1)$ 番目の分析値までが次のグループとして分割され、このグループに対し、新たな特徴量 A_2 が得られる。このような処理を繰り返して行ない、1つのフレームの $(k-q)$ 番目の分析値から k 番目の分析値までが最後のグループとして分割され、このグループに対して、新たな特徴量 $A_{(k-q)}$ が得られる。 $A_{(k-q)}$ が得られた後、 i は $(k-q)$ よりも大きくなるので（ステップS27）、1つのフレームについてのグループ分割処理、並びに演算処理を終了する。

【0049】次いで、次のフレームについての同様のグループ分割処理並びに演算処理を行ない、 M 番目のフレーム（最後のフレーム）に対して処理がなされると、全ての処理を終了する。

【0050】以上のようにして新たな特徴パターン A_i を得ることができる。図11は1つのフレームについてなされる上記グループ分割処理並びに演算処理の具体例を示す図である。この例では、バンドパスフィルタ群1が15個のバンドパスフィルタからなり、15個のバンドパスフィルタによって1フレーム当り分析次元 k が

(6)

特開平5-19782

9

“15”の分析値（すなわち15個の分析値）が得られたとした場合が示されている。この際に、1つのグループ内の分析値の個数 q を“5”にすると、1つのフレームでは、先づ、第1番目の分析値 f_1 から第5番目の分析値 f_5 までが最初のグループとして分割され、このグループ内の各分析値 $f_1 \sim f_5$ の加算演算がなされて A_1 が求まる。次いで、分析値を分析軸方向に1つずらす。これにより、第2番目の分析値 f_2 から第6番目の分析値 f_6 までが次のグループとして分割され、このグループ内の各分析値 $f_2 \sim f_6$ の加算演算がなされて A_2 が求まる。このようにして、分析値を1つづつずらし分析値を一部重複させながら（いまの場合、4個の分析値を重複させながら）グループ分割し、1つのフレームにおいて、11個に分割されたグループで11次元の新しい特徴ベクトル $A_1 \sim A_{11}$ を得ることができる。このようにして得られた11次元の新しい特徴ベクトル $A_1 \sim A_{11}$ は、グループ間にまたがる音声の分析軸方向の微小変動を吸収したものとなっており、この特徴ベクトル $A_1 \sim A_{11}$ を用いてマッチングを行ない認識結果を得ることができる。

【0051】なお、上述の例では、1つのグループ内の分析値の数 q を“5”としたが、必ずしも“5”である必要がなく、任意の個数に設定できる。また、分析値を1つづつずらしてグループ分割したが、例えば2つづつずらしてグループ分割しても良い。

【0052】ところで、上述した新しい特徴ベクトル $A_1 \sim A_{11}$ は、音声の分析軸方向の微小変動を吸収したものとはいるが、時系列方向の微小変動に対しては対処がなされていない。すなわち、微小変動は、分析軸方向だけではなく時系列方向にも発生するので、分析軸方向のみならず、時系列方向についても対処がなされるのが良い。

【0053】このような処理は、未知入力音声の特徴ベクトルに変換し変換された特徴ベクトルを用いてマッチングを行ない認識結果を得る過程において、音響分析後の所定の数フレームを分割すべきグループの対象にしてから、例えば第3の実施例で述べたと同様の処理操作をすることによって行なうことができる。なお、この場合の装置構成は、図9に示したものに对应したものを使用することができる。

【0054】図12は分析軸方向のみならず時系列方向についても音声の微小変動を吸収する流れを示すフローチャートである。図12において、全体の処理流れステップS31～S37は、図10に示した処理流れS21からS27と対応したものとなっているが、図12では、ステップS33において、 $(p+r)$ 番目のフレームまでの r 個の複数フレームにわたって加算処理がなされる。すなわち、 $q \times r$ の大きさの窓に含まれる分析値を1つのグループとして抽出し、このグループ内の $q \times r$ 個の分析値 f_{lk} の加算演算を行なうことにより、1つ

10

の新しい特徴ベクトル A_{p1} を得ることができる。

【0055】図13は上記処理の具体例を示す図であり、この例では、1フレーム当りの分析次元 k が“15”，1つのグループ内の分析値の個数 q が“3”，1回の処理におけるフレーム数 r が“3”の場合が示されており、この場合には、第 p 番目のフレームから第 $(p+2)$ 番目のフレームにわたって分析値を1つづつ分析軸方向にずらしながらグループ分割をし、各グループでステップS33の加算演算を行なって、13個の特徴ベクトル $A_{p1} \sim A_{p13}$ を得ることができ、これにより、時系列方向の微小変動をも吸収することができた。

【0056】なお、図10、図12に示した処理によって得られた特徴ベクトルを認識処理にそのまま用いるときには、第1の実施例で説明したのと同じ問題が生じる可能性がある。従って、この場合にも第2の実施例におけると同様に正規化処理を行なうのが良い。

【0057】図14は本発明の第4の実施例のブロック図であって、この第4の実施例では、第2の実施例と同様な構成にし、第3の実施例における問題を回避するようにしている。すなわち、この第4の実施例の特徴抽出部21では、演算部22が、グループ値計数部23と、正規化部24とから構成され、第3の実施例の演算部18における演算処理機能、すなわち図10または図12の処理に加えて、さらに正規化処理機能を有している。

【0058】次にこのような第4の実施例の音声特徴抽出装置の処理動作を図15のフローチャートを用いて説明する。先づ、ステップS41では、フレームの番号 p を“0”に初期設定する。次いで、第3の実施例の演算部18と同様に、図10または図12のフローチャートに示す処理を行ない（ステップS42）、特徴ベクトル A_{lx} を得る。次いで、ステップS43で正規化処理を行なう。

【0059】この正規化処理は、図10または図12の処理によって得られた各特徴ベクトル A_{lx} を r フレーム内の特徴ベクトル A_{lx} の合計値 S で除算することによってなされ、この結果、最終的な特徴量 $B(p+1)h$ を求めることができる。

【0060】しかる後、ステップS44において p を“1”だけ歩進し、再びステップS42に戻る。この結果、窓は、時系列方向に“1”だけずらされ、ずらされた窓に含まれる各特徴量 A_{lx} に対しても、上記と同様の正規化処理がなされる。このようにして、第2の実施例と同様の正規化を行なうことができ、ステップS45において p が $(M-r)$ となると、全ての演算処理を終了する。

【0061】これによって、第2の実施例と同様に、図4(a)，(b)のような特徴ベクトルを、最終的に図8(a)，(b)のような正規化された特徴ベクトルにそれぞれ変換することができる。

【0062】このように、第4の実施例では、音響分析

(7)

特開平5-19782

11

された結果の分析値に対し、分析軸方向の微小変動あるいは、分析軸方向並びに時系列方向の微小変動を吸収した新しい特徴パターンを求め、さらにその上で、正規化処理を施すことによって、音量の大小差を吸収することができた。すなわち、第2、第4の実施例は、全分析値に対して各分析値がどれ位の割合を占めているかを計算する正規化処理がなされ、分析軸、時系列、あるいはその両方に沿った値の変化の割合を同じにすることができるので、これによって、音量の大小差を吸収することができた。

【0063】また、上述した第1～第4の各実施例では、音声の微小変動を吸収した特徴ベクトルを作成可能となり、特に、第3、第4の実施例では、グループ間にまたがった微小変動をも吸収した特徴ベクトルを作成可能となって、これにより認識精度を向上させることができるが、上記のように作成された特徴ベクトルは、時系列フレームの瞬時の特徴を表わしている。一方、近年において、該当する特徴ベクトルに対し前後フレームの特徴ベクトルの“動き”をも考慮した特徴（以後、動的特徴と呼ぶ）が提案されている。この動的特徴を用いた例として、△ケプストラム法による音声認識が極めて有効であることが種々の実験で実証され、現在、音声認識系で広く用いられている。なお、△ケプストラム法とは、例えば文献「階層的スペクトル動特性を用いた音声認識 信学技報、SP90-60、第25～31頁、1990年12月」に示されているように、簡単には、ケプストラムや対数パワーの線形回帰係数を求め、これらの時系列を特徴ベクトルとして音声認識を行なう方法である。

【0064】しかしながら、△ケプストラム法を用いる場合には、特徴を抽出するのに、回帰係数を求めなければならないので、動的特徴に簡単に变换することができないという欠点がある。

【0065】図16は本発明の第5の実施例のブロック図であり、この第5の実施例では、△ケプストラム法によらずに、動的特徴に变换することを意図している。なお、図16において、図1、図5、図9、図14と対応する箇所には同じ符号を用いている。図16を参照すると、第5の実施例では、特徴抽出部30は、バンドパスフィルタ群1と、区間検出部2と、時系列方向に所定のフレーム数からなるフレーム枠を設けるフレーム枠設定部31と、フレーム枠を時系列方向に走査することによって1つの特徴ベクトルを得るフレーム枠走査部32とを有している。

【0066】次にこのような構成の第5の実施例の音声特徴抽出装置の動作を図17のフローチャートを用いて説明する。なお、以下では、音響分析次元数を k 、フレーム枠数を $M1$ 、フレーム走査数を $M2$ とする。また、フレーム p 内の第 k 番目の分析値を f_{pk} で表わし、フレーム p 内の j 番目の特徴ベクトルを α_{pj} で表わしてい

12

る。

【0067】先づ、ステップS51、S52、S53では、フレーム走査番号 $m2$ 、フレーム枠番号 $m1$ 、1フレーム内の分析値の番号 k を“0”に初期化する。次いで、ステップS54～S58では、フレーム $(p+1)$ から数えて $M1$ 番目のフレームまでのフレームに含まれる分析値を取り出す。しかる後、ステップS59では、フレーム走査番号 $m2$ を“1”歩進ませてフレーム枠を時系列方向に1フレーム走査し、再びステップS52に戻って、その枠内に含まれる分析値を同様にして取り出す。この操作をステップS60において $m2$ が $M2$ となるまで、すなわち $M2$ 走査するまで繰り返す。このとき、1つの特徴ベクトルは、いままで取り出した全分析値となる。

【0068】図18は上記処理の具体例を示す図である。この例では、バンドパスフィルタ群1が15個のバンドパスフィルタからなり、15個のバンドパスフィルタによって1フレーム当り分析次元が“15”の分析値（すなわち15個の分析値）が得られたとした場合が示されている。この際に、図15の左側に示すように、フレーム枠を“3”にし、走査数を“3”にした場合、1回目のフレーム枠の設定で、同図の右側に示すように、特徴ベクトルの1～45次元目（A1～C15）までを作成する。同様の走査による2回目、3回目のフレーム枠の設定で特徴ベクトルB1～D15、C1～E15を作成していき、最終的に、135次元の1つの特徴ベクトルに変換することができる。しかる後、この特徴ベクトルは、辞書内に予め同様の処理によって登録されている標準パターンとDPマッチングによる照合がなされ、認識結果が出力される。

【0069】このように第5の実施例では、従来の△ケプストラム法のように、特徴ベクトルそのものが動的特徴を隔に表わしているわけではないが、すでに存在する特徴ベクトルの数フレームをまとめて1つの特徴ベクトルにすることによって、間接的に動的特徴量に変換することができることになる。

【0070】ところで、上述のように動的特徴量を作成することができるが、図18に示したような特徴ベクトルは、音声の微小変動を吸収したものとはなっていない。

【0071】図19は本発明の第6の実施例のブロック図である。この第6の実施例では、上記第5の実施例の問題点を解決することを意図したものであって、未知入力音声の特徴量に変換し、変換された特徴量を用いてマッチングを行ない認識結果を得る過程において、音響分析された結果得られる各分析値をある次数ごとにグループにまとめ、第5の実施例の操作をすることによってグループ数を増やし、各グループ内の分析値を計数することによって特徴量に変換するようにしている。

【0072】このため、第6の実施例では、図9に示す

(8)

特開平5-19782

13

ように、特徴抽出部33は、時系列方向に所定のフレーム数からなるフレーム枠を設けるフレーム枠設定部34と、分析値をある回数ごとにグループにまとめるグループ化部35と、グループ内の分析値に所定の演算処理（加算処理）を施す演算部36と、フレーム枠を時系列方向に走査することによって1つの特徴ベクトルを得るフレーム枠走査部37とを有している。なお、演算部36は、グループ値計数部38により構成されている。

【0073】次にこのような構成の第6の実施例の音声特徴抽出装置の動作を図20のフローチャートを用いて説明する。なお、以下では、1グループ内の分析値数をGとする。先づ、ステップS61、S62では、フレーム走査番号m2、1フレーム内の分析値の番号kを“0”に初期化する。次いで、ステップS63～S65では、フレームpの分析値fをグループごとに合計する。ここで、グループとは、分析値をフレーム及び分析軸の数次元ごとにまとめ上げたものをいう。次にグループ内の分析値を全て合計する。しかる後、ステップS66では、第5の実施例と同様、フレーム走査番号m2を“1”歩進ませてフレーム枠を時系列方向に1フレーム走査し、再びステップS62に戻って、同様の処理を繰り返す。この操作をステップS67においてm2がM2となるまで、すなわちM2走査するまで繰り返し、1つの特徴ベクトルを作成する。

【0074】図21は上記処理の具体例を示す図である。図21は図18と対応したものとなっているが、図21では、フレーム枠を設定し、グループ化を行なっている。すなわち、図21の左側に示すように、フレーム枠数“3”、分析次元数“5”を1つのグループとし、“3×5”のグループ内の分析値を加算演算して特徴ベクトルを求めている。具体的には、1回目のフレーム枠の設定で、同図右側に示すように特徴ベクトルの1～3次元目（ $\alpha 1 \sim \alpha 3$ ）までを作成する。すなわち、A1からC5までの15個の分析値を最初のグループとして、 $\alpha 1$ を作成し、A6～C10までの15個の分析値を次のグループとして、 $\alpha 2$ を作成し、A11～C15までの15個の分析値を最後のグループとして、 $\alpha 3$ を作成する。次いで、2回目のフレーム枠の設定で、特徴ベクトル $\alpha 4 \sim \alpha 6$ を作成し、3回目のフレーム枠の設定で、特徴ベクトル $\alpha 7 \sim \alpha 9$ を作成する。このようにして、図21の例では、9次元の1つの特徴ベクトルが作成される。なお、この例では、1つのフレーム枠当りのグループ数を“3”にし、1グループ内の分析軸方向の分析値数を“5”にしたが、これを増減し、グループ数を増減することも可能である。

【0075】第6の実施例の操作によって、定められたグループ内での音声の微小変動を収集することができた。しかしながら、この第6の実施例におけるグループ分割では、グループ間で共有する要素、すなわち共有の分析値がないため、グループ間にまたがって微小変動が

14

生じた場合には、その変動を吸収できない。このような問題を解決するためには、第1、第2の実施例を改良して第3、第4の実施例としたのと同様の仕方で、グループ間で分析値が一部重複するようにグループ分割をするのが良い。

【0076】図22は第7の実施例のブロック図であり、第7の実施例では、上記第6の実施例を改良し、グループ間で分析値が一部重複するようにグループ分割することを意図している。すなわち、第7の実施例では、図22に示すように、特徴抽出部39は、図19に対応した構成となっているが、特徴抽出部39のグループ化部40は、各分析値を分析軸方向に一部重複させながらグループ分割する点で、第6の実施例と相違している。すなわち、第7の実施例では、各分析値を例えば1つづつずらしながらグループ分割するようにしており、この結果、第6の実施例に比べて、グループ数を増加させることができる。また、特徴抽出部39の演算部41のグループ値計数部42は、上記のように一部重複させて分割されたグループ内の要素、すなわち分析値の和を計数するようになっている。

【0077】次にこのような構成における第7の実施例の音声特徴抽出装置の動作を図23のフローチャートを用いて説明する。先づ、ステップS71、S72では、フレーム走査番号m2、1フレーム内の分析値の番号kを“0”に初期化する。次いで、ステップS73～S75では、グループ内の分析値を合計した後、グループの対象を分析軸方向に1つづつずらす。しかる後、ステップS76では、第6の実施例と同様、フレーム走査番号m2を“1”歩進ませてフレーム枠を時系列方向に1フレーム走査し、再びステップS72に戻って、同様の処理を繰り返す。この操作をステップS77においてm2がM2となるまで、すなわちM2走査するまで繰り返し、1つの特徴ベクトルを作成する。

【0078】図24は上記処理の具体例を示す図である。この例では、フレーム枠を設定し、グループ化を行なう際に、図24の左側に示すように、フレーム枠数“3”、分析次元数“5”を1つのグループとし、“3×5”のグループ内の分析値を加算演算して特徴ベクトルを求め、次に、グループの対象を分析軸方向に1つづつシフトし、同様にグループ内の分析値を加算演算して特徴ベクトルを求める。具体的には、1回目のフレーム枠の設定で、同図右側に示すように特徴ベクトルの1～11次元目（ $\alpha 1 \sim \alpha 11$ ）までを作成する。すなわち、A1からC5までの15個の分析値を最初のグループとして、 $\alpha 1$ を作成し、A2～C6までの15個の分析値を次のグループとして、 $\alpha 2$ を作成し、これを順次繰り返し、A11～C15までの15個の分析値を最後のグループとして、 $\alpha 11$ を作成する。次いで、2回目のフレーム枠の設定で、特徴ベクトル $\alpha 12 \sim \alpha 22$ を作成し、3回目のフレーム枠の設定で、特徴ベクトル $\alpha 23 \sim \alpha$

(9)

特開平5-19782

15

33を作成する。このようにして、図24の例では、3次元の1つの特徴ベクトルが作成される。なお、この例では、1つのフレーム枠当りのグループ数を“3”にし、1グループ内の分析軸方向の分析値数を“5”にしたが、これを増減し、グループ数を増減することも可能である。

【0079】第7の実施例の操作によって、定められたグループ間にまたがって音声の微小変動が生じる場合にも、これを吸収することができた。

【0080】このように、第5乃至第7の実施例では、基本的に、時系列方向に所定のフレーム数からなるフレーム枠を設定し、フレーム枠を時系列方向に走査することによって1つの特徴ベクトルを得るようにしており、これにより、△ケプストラム法によらずに、動的特徴に簡単に変換することができた。

【0081】しかしながら、第5乃至第7の各実施例では、フレーム枠の走査をしているため、最終的に得られる1つの特徴ベクトルの次元数は、音響分析結果である分析値の分析次元数よりも通常増加してしまう。従って、音声認識時において特徴ベクトルと辞書とのマッチングの計算量は、音響分析された結果の分析値をそのまま用いてマッチングする場合の計算量よりも増加する。

【0082】図25は本発明の第8の実施例のブロック図であって、この第8の実施例では、特徴ベクトルと辞書とのマッチングの計算量を削減するために、各フレーム枠単位で得られた特徴ベクトルを加算するようにしている。すなわち、図25を参照すると、第8の実施例では、特徴抽出部43は、時系列方向に所定のフレーム数からなるフレーム枠を設けるフレーム枠設定部31と、今回のフレーム枠設定時作成した特徴ベクトルを前回のフレーム枠設定時に作成された特徴ベクトルと加算する特徴ベクトル加算部44と、フレーム枠を時系列方向に走査するフレーム枠走査部32とを有している。なお、図25は図16の構成（すなわち第5の実施例の構成）を改良したものとして構成されており、図16と同様の箇所には同じ符号を付している。

【0083】次にこのような構成の第8の実施例の音声特徴抽出装置の動作を図26のフローチャートを用いて説明する。先づ、ステップS81、S82、S83では、フレーム走査番号m2、フレーム枠番号m1、1フレーム内の分析値の番号kを“0”に初期化する。次いで、ステップS84～S88では、前回フレーム枠の設定時に作成した特徴ベクトルと今回のフレーム枠設定時に作成した特徴ベクトルとを加算する。しかる後、ステップS89では、フレーム走査番号m2を“1”歩進ませてフレーム枠を時系列方向に1フレーム走査し、再びステップS82に戻って、同様の処理を繰り返す。この操作をステップS90においてm2がM2となるまで、すなわちM2走査するまで繰り返し、1つの特徴ベクトルを作成する。

16

【0084】図27は上記処理の具体例を示す図である。図18に示したような処理により、図27の左側に示すように、1回目のフレーム枠の設定で、45次元の特徴ベクトル(A1～C15)が作成され、2回目のフレーム枠の設定で、45次元の特徴ベクトル(B1～D15)が作成され、3回目のフレーム枠の設定で、45次元の特徴ベクトル(C1～E15)が作成されたとする。

【0085】この場合、第8の実施例では、1回目、2回目、3回目のフレーム枠設定でそれぞれ得られた45次元の各特徴ベクトル(A1～C15)、(B1～D15)、(C1～E15)を互いに加算し、図27の右側に示すように、合計で45次元の1つの特徴ベクトル($\alpha 1 \sim \alpha 45$)を作成する。これにより、第5の実施例と比べた場合、1つの特徴ベクトルの次元数を1/3に削減することができ、この結果、音声認識時において特徴ベクトルと辞書とのマッチングの計算量を削減することができた。

【0086】なお、上記例では、第5の実施例を改良したものとして構成し、第5の実施例により作成された特徴ベクトルを用いたが、第6、第7の実施例についても同様にして改良することができ、第6、第7の実施例により作成された特徴ベクトルに対しても同様の処理によって1つの特徴ベクトルの次元数を削減することができる。

【0087】さらに、前述の第1の実施例、第3の実施例で述べたと同様に、第5乃至第8の実施例によって作成された特徴ベクトルをそのまま用いる場合には、音声の大小などにより分析値が変化し、分析軸における分布の形状が同一カテゴリに分類されるべきパターンであっても値が異なるために、正しく分類できなくなる可能性が生ずる。例えば、同一の音声を発声しても音量の大小により周波数変換された値は、図4(a)、(b)に示したようになり、それぞれ異なったカテゴリのものに分類されてしまう恐れがある。

【0088】このような問題を解決するため、第1、第3の実施例に対し第2、第4の実施例の構成にしたと同様に、第5乃至第8の実施例に対しても以下のような構成にすることができる。

【0089】図28は本発明の第9の実施例のブロック図であり、図28の例では、第8の実施例に対し改良がなされた場合が示されている。すなわち、この第9の実施例の特徴抽出部46は、図25に示した特徴抽出部43に対し、さらに正規化部47が付加されている。

【0090】次にこのような構成における第9の実施例の音声特徴抽出装置の処理動作を図29のフローチャートを用いて説明する。先づ、ステップS91ではフレームの番号pを“0”に初期設定し、次いで、ステップS92で図26のフローチャートに示した処理を行ない、しかる後、ステップS93～S95で第2、第4の実施例で説明したと同様の正規化処理を行なう。

(10)

特開平5-19782

17

【0091】すなわち、この正規化処理では、図26に示した処理によって得られた特徴ベクトル α を正規化して特徴ベクトル β を求める。なお、正規化操作は、各特徴ベクトル α をその合計値 S で除算することにより行なわれ、この処理を時系列方向に順次にずらしながら最後のフレームとなるまで繰り返し行なう。

【0092】上記の例では、図26のフローチャートに示した処理の結果に対して正規化処理を施した場合を示したが、図17、図20、図23のフローチャートに示した処理の結果に対しても同様の手順で正規化処理を行なうことができる。

【0093】例えば、図17のフローチャートに示した処理の結果に対して正規化処理を施す場合、図18の右側に示した135次元の1つの特徴ベクトルに対して正規化処理がなされる。この結果、図4(a)、(b)のような特徴ベクトルを最終的に図8(a)、(b)のような正規化された特徴ベクトルに変換することができる。音量の大小差をも吸収することができる。

【0094】なお、上述した各実施例では、音響分析を行なうのに、バンドパスフィルタ群1を用いているが、これにかわって例えばFFT等の他の手段を用いることもできる。さらに、音響分析として周波数分析を用いているが、これに限らず、LPCなどによって音響分析を行なっても良い。

【0095】また、認識部6には、DPマッチングを用いて説明したが、ニューラルネットワーク、HMMなどにより識別を行なわせるようにしても良い。

【0096】

【発明の効果】以上に説明したように、請求項1記載の発明によれば、入力音声を音響分析した結果の各分析値をグループに分割し、分割された各グループ内の分析値に所定の演算処理を施して特徴量を生成するようにしているため、音声の微小変動を吸収することができる。

【0097】また、請求項2記載の発明によれば、分割された各グループは、互いに隣接したグループ間において少なくとも一部が重複している分析値を有しているため、グループ間にまたがる音声の微小変動をも吸収することができる。

【0098】また、請求項3記載の発明によれば、分析軸方向に分割されるようになっているため、分析軸方向の微小変動を吸収することができる。

【0099】また、請求項4記載の発明によれば、時系列方向に分割されるようになっているため、時系列方向の微小変動を吸収することができる。

【0100】また、請求項5記載の発明によれば、各グループ内の分析値を加算し特徴量を生成するようになっているため、特徴量を削減することができ、辞書とのマッチングにおける計算量を減らすことができる。

【0101】また、請求項6記載の発明によれば、各グループ内の分析値に所定の演算処理を施して生成された

18

特徴量に対し正規化処理を施し、最終的な特徴ベクトルを生成するようになっているため、音量の大小差を吸収することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施例のブロック図である。

【図2】第1の実施例の音声特徴抽出装置の動作を説明するためのフローチャートである。

【図3】第1の実施例における演算処理の具体例を示す図である。

【図4】(a)、(b)は分析軸に沿った分布の形状が互いに同じであるが、各値が互いに異なっている特徴ベクトルの一例を示す図である。

【図5】本発明の第2の実施例のブロック図である。

【図6】第2の実施例の音声特徴抽出装置の動作を説明するためのフローチャートである。

【図7】第2の実施例における演算処理の具体例を示す図である。

【図8】(a)、(b)は図4(a)、(b)に示す特徴ベクトルを正規化した結果をそれぞれ示す図である。

【図9】本発明の第3の実施例のブロック図である。

【図10】第3の実施例の音声特徴抽出装置の動作を説明するためのフローチャートである。

【図11】第3の実施例における処理の具体例を示す図である。

【図12】分析軸方向のみならず時系列方向についても音声の微小変動を吸収する処理の流れを示すフローチャートである。

【図13】図12の処理の具体例を示す図である。

【図14】本発明の第4の実施例のブロック図である。

【図15】第4の実施例の音声特徴抽出装置の動作を説明するためのフローチャートである。

【図16】本発明の第5の実施例のブロック図である。

【図17】第5の実施例の音声特徴抽出装置の動作を説明するためのフローチャートである。

【図18】第5の実施例における演算処理の具体例を示す図である。

【図19】本発明の第6の実施例のブロック図である。

【図20】第6の実施例の音声特徴抽出装置の動作を説明するためのフローチャートである。

【図21】第6の実施例における演算処理の具体例を示す図である。

【図22】本発明の第7の実施例のブロック図である。

【図23】第7の実施例の音声特徴抽出装置の動作を説明するためのフローチャートである。

【図24】第7の実施例における演算処理の具体例を示す図である。

【図25】本発明の第8の実施例のブロック図である。

【図26】第8の実施例の音声特徴抽出装置の動作を説明するためのフローチャートである。

【図27】第8の実施例における演算処理の具体例を示

(11)

特開平5-19782

19

20

す図である。

【図28】本発明の第9の実施例のブロック図である。

【図29】第9の実施例の音声特徴抽出装置の動作を説明するためのフローチャートである。

【図30】従来の一般的な音声認識システムの構成図である。

【図31】(a), (b)は入力音声が微小変動したときの周波数分析結果を説明するための図である。

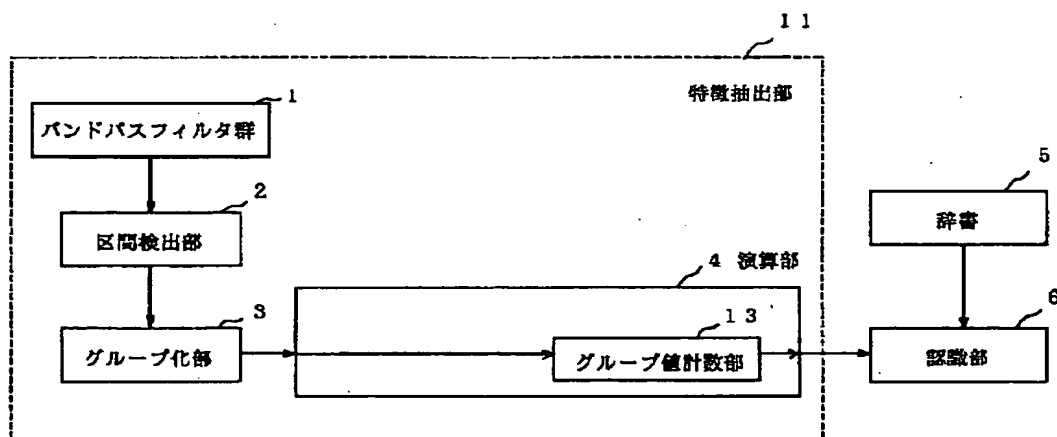
【図32】従来の特徴ベクトルの作成の仕方を説明するための図である。

【符号の説明】

1
バンドパスフィルタ群
2
区間検出部
3, 17, 35, 40
グループ化部
4, 14, 18, 36, 41

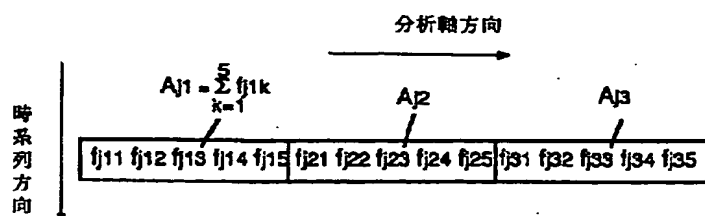
演算部
5
辞書
6
認識部
11, 12, 16, 21, 30, 33, 39, 43
特徴抽出部
13, 19, 23, 38, 42
グループ値計数部
10 15, 24, 47
正規化部
31, 34
フレーム枠設定部
32, 37
フレーム枠走査部
44
特徴ベクトル加算部

【図1】



【図3】

【図8】



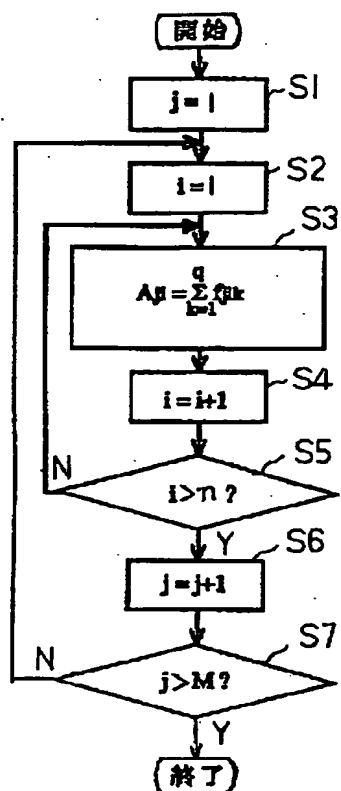
(a) 0.1 0.2 0.2 0.1 0.05 0.05 0.1 0.05 0.05

(b) 0.1 0.2 0.2 0.1 0.05 0.05 0.1 0.05 0.05

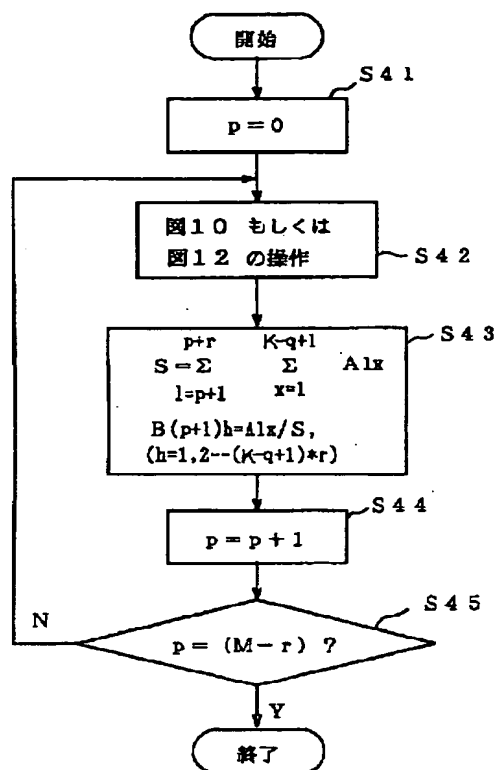
(12)

特開平5-19782

【図2】



【図15】

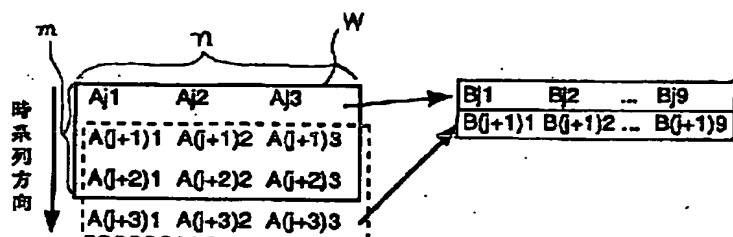


【図4】

(a) 30 60 60 30 15 15 30 15 15

(b) 60 120 120 60 30 30 60 30 30

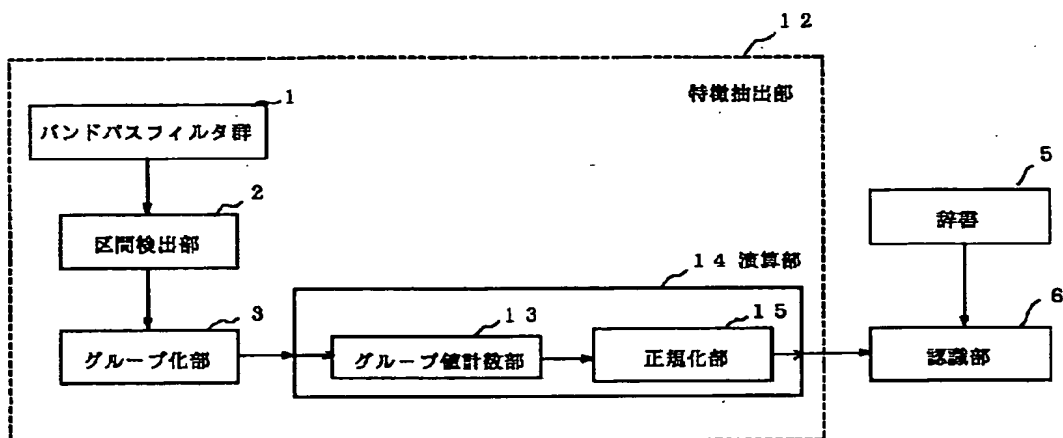
【図7】



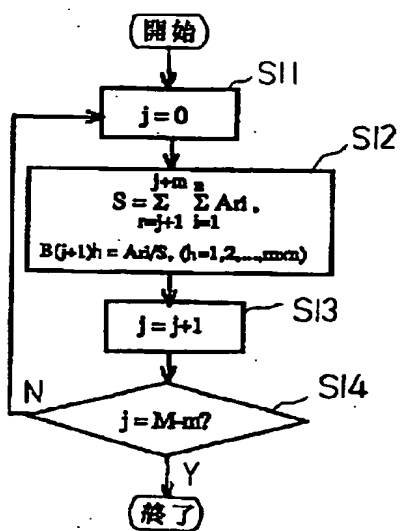
(13)

特開平5-19782

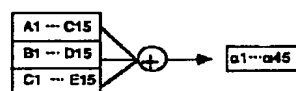
【図5】



【図6】



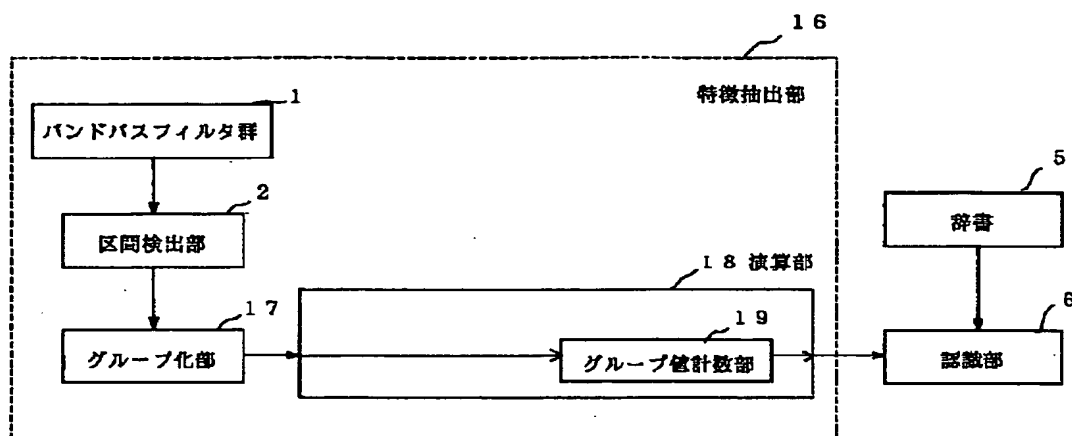
【図27】



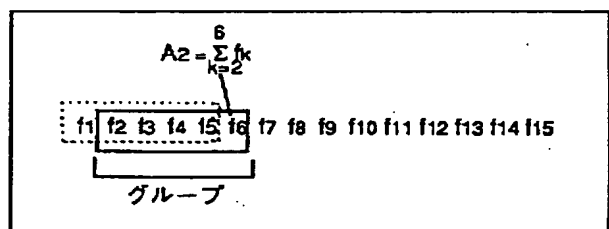
(14)

特開平5-19782

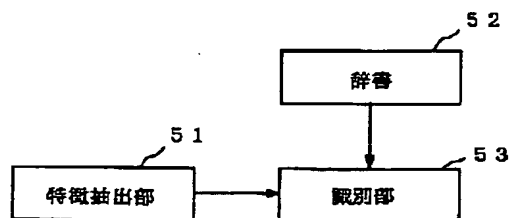
【図9】



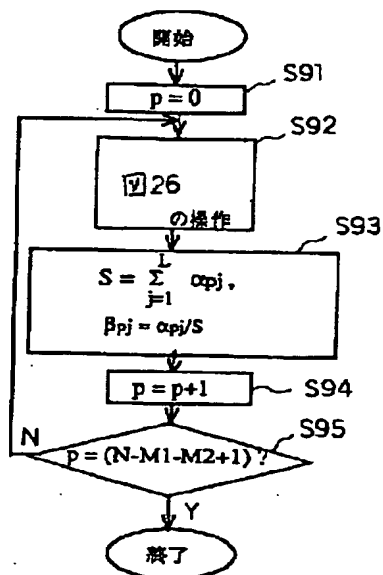
【図11】



【図30】



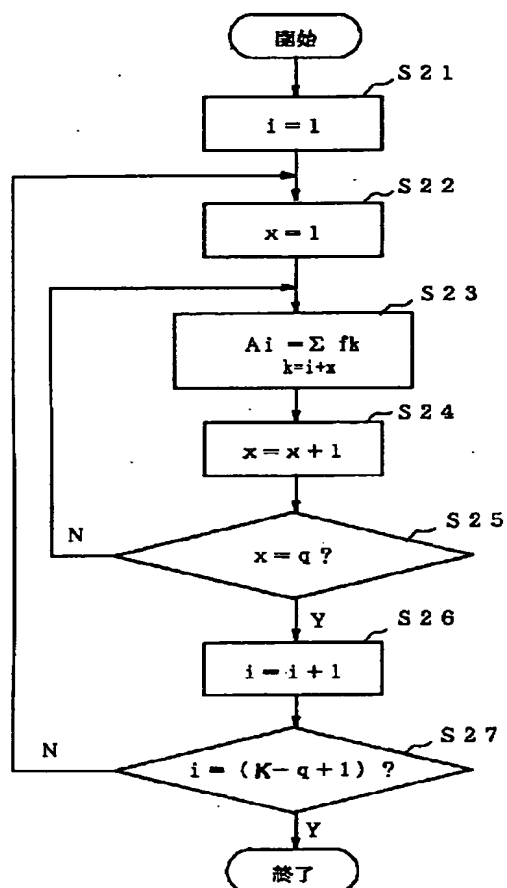
【図29】



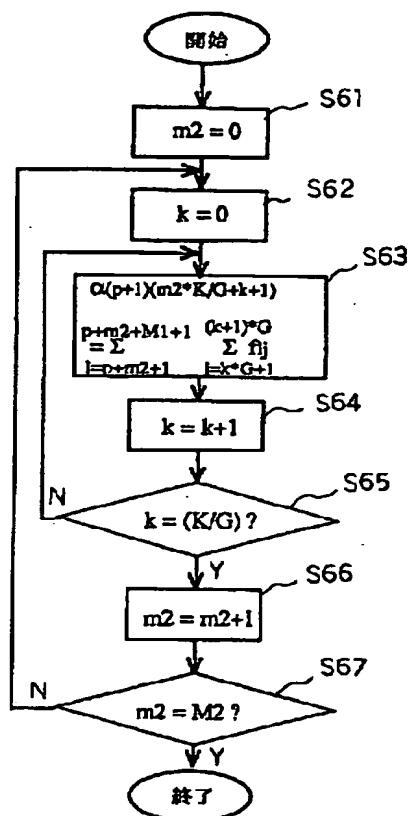
(15)

特開平 5 - 1 9 7 8 2

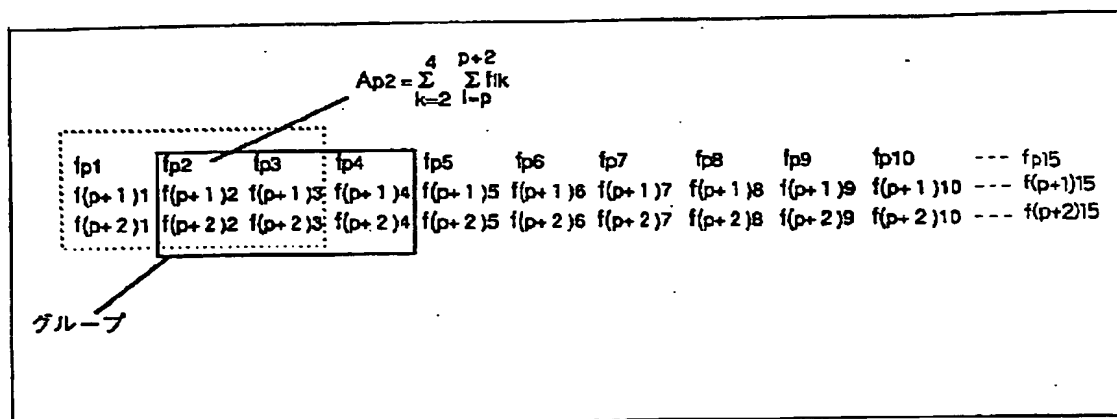
【図 10】



【図 20】



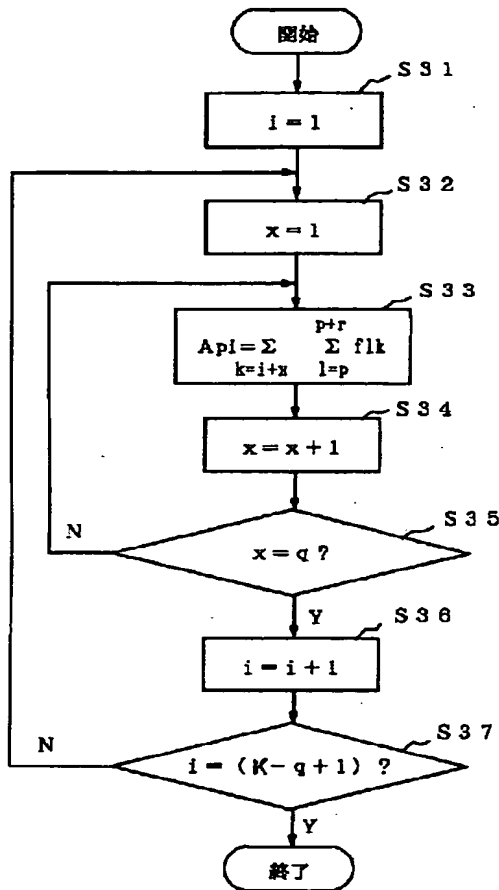
【図 13】



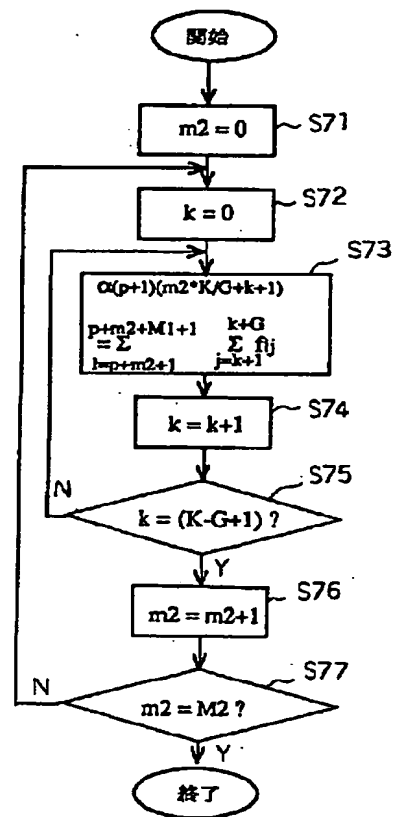
(16)

特開平5-19782

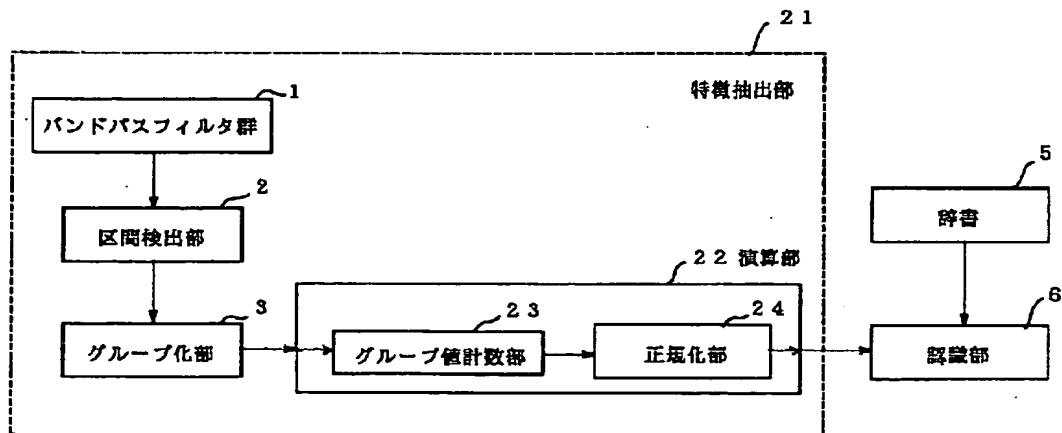
【図12】



【図23】



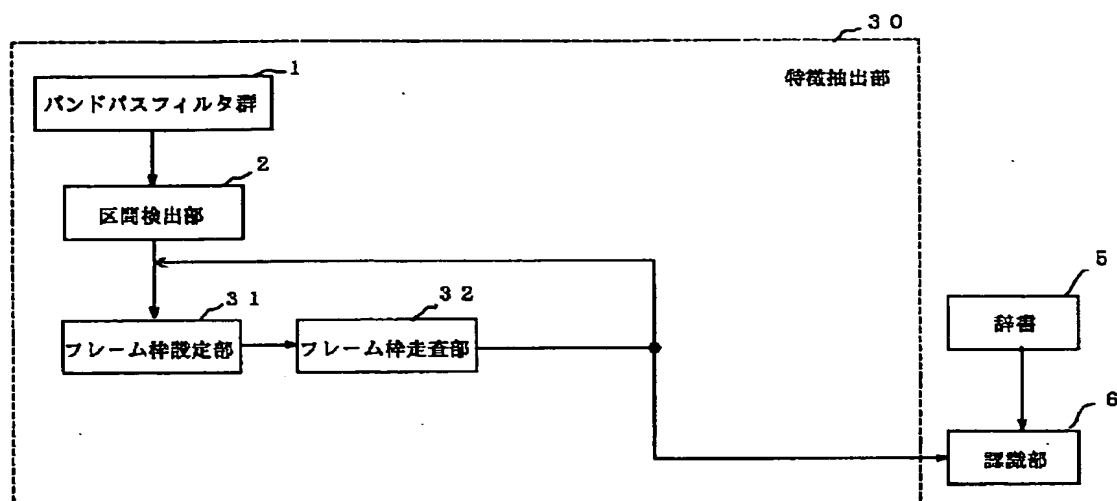
【図14】



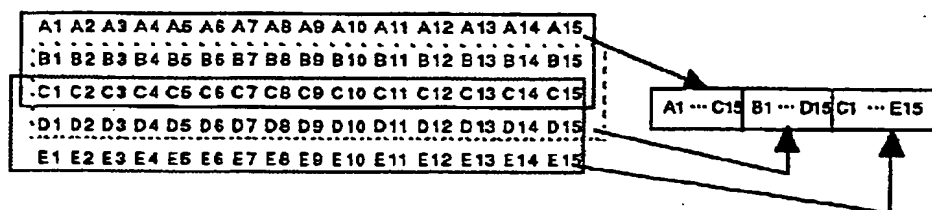
(17)

特開平5-19782

【図16】



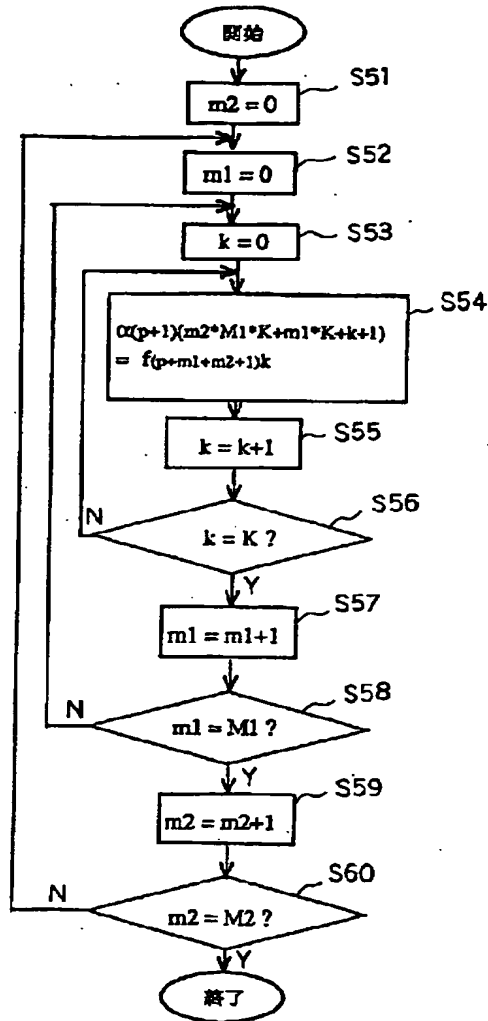
【図18】



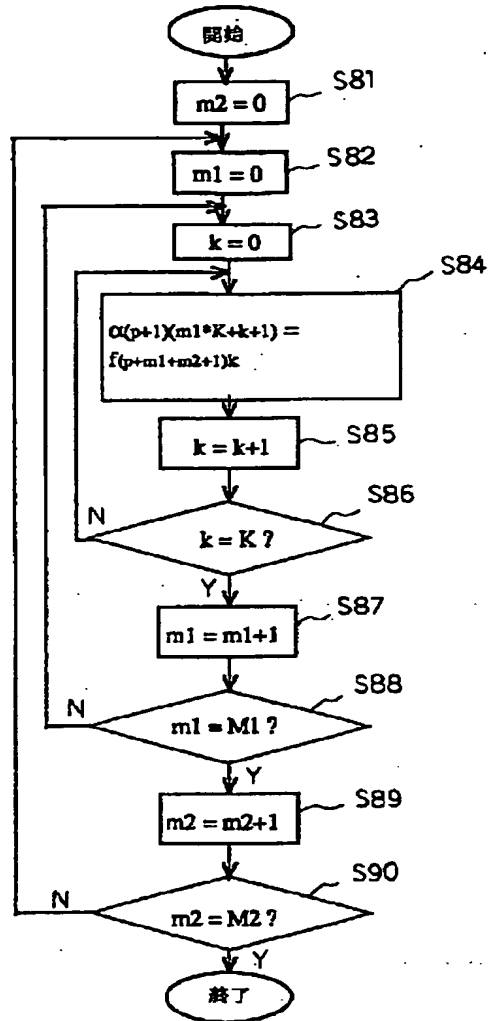
(18)

特開平5-19782

【図17】



【図26】



【図31】

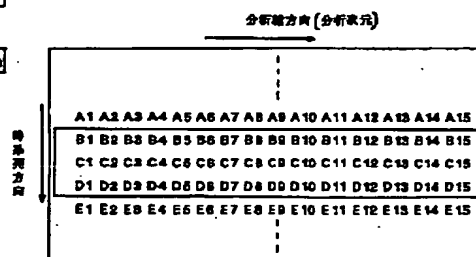
(a)

6	18	12	14	0	0	0	8	8	0	0	2	0	14	0
---	----	----	----	---	---	---	---	---	---	---	---	---	----	---

(b)

6	17	9	13	5	0	1	5	5	1	0	0	3	18	0
---	----	---	----	---	---	---	---	---	---	---	---	---	----	---

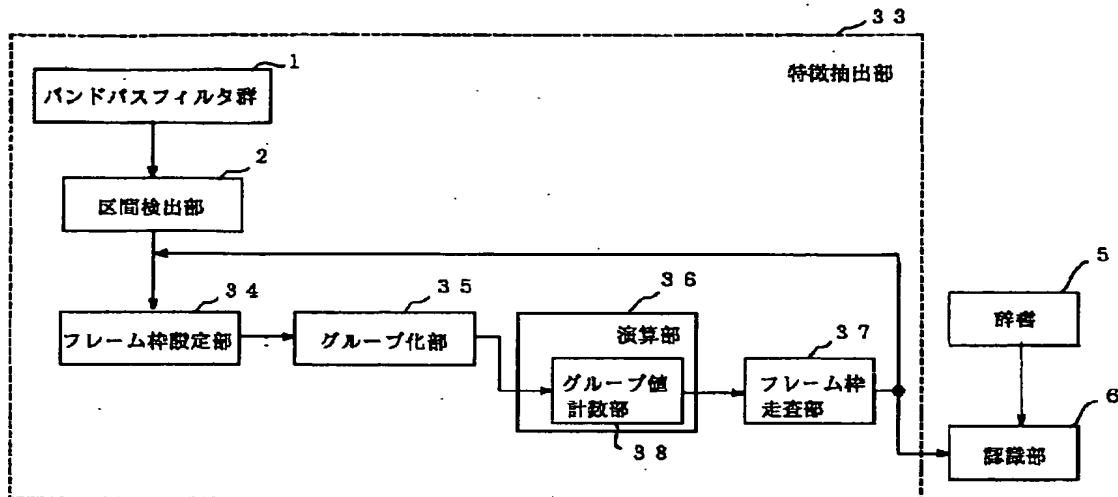
【図32】



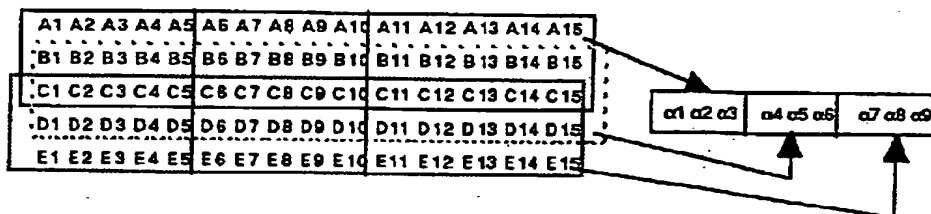
(19)

特開平5-19782

【図19】



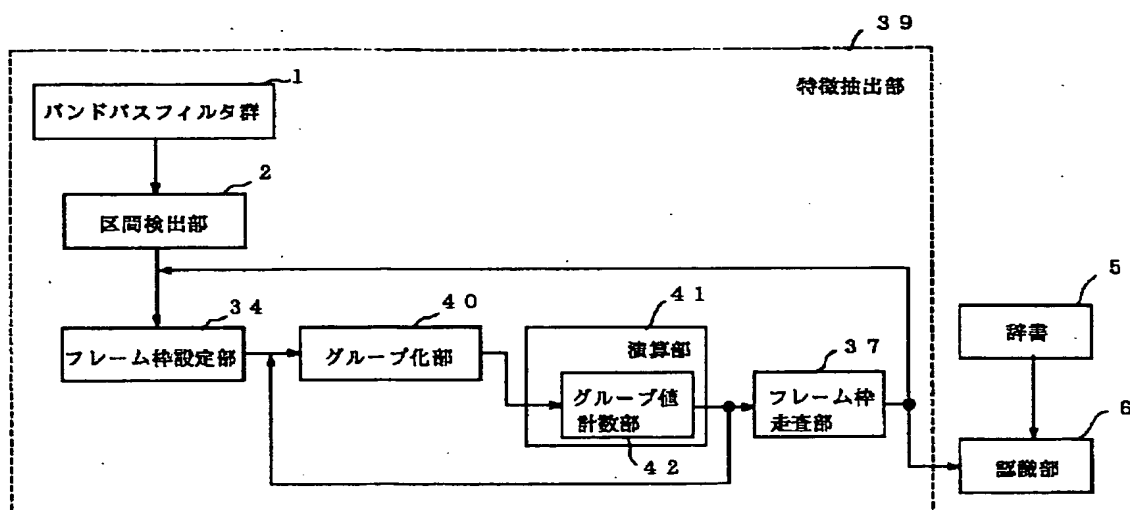
【図21】



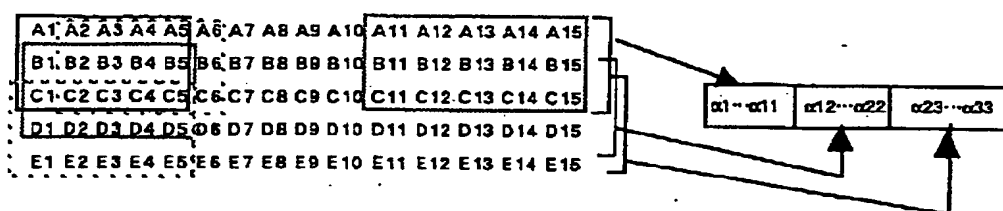
(20)

特開平5-19782

【図22】



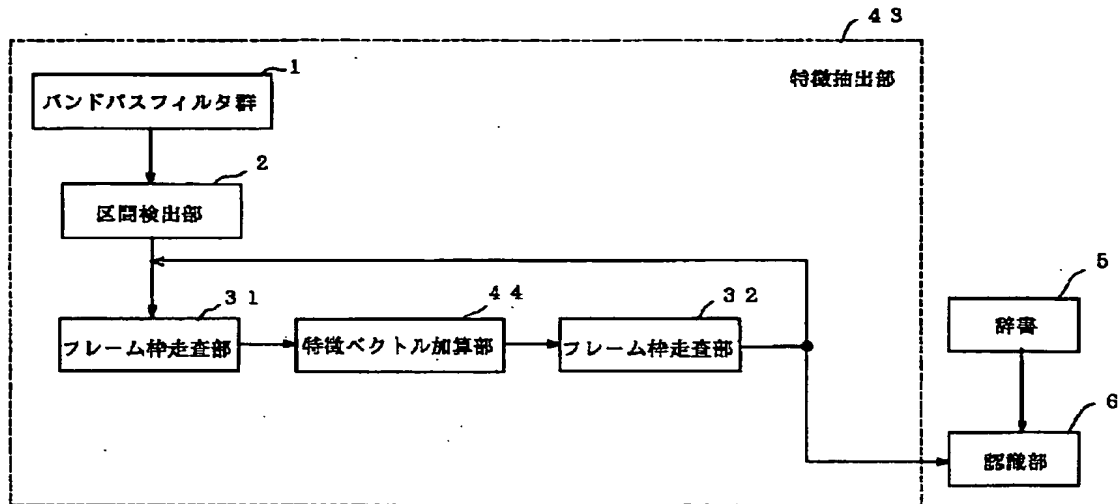
【図24】



(21)

特開平5-19782

【図25】



【図28】

